International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:15, No:8, 2021

Classification of Extreme Ground-Level Ozone Based on Generalized Extreme Value Model for Air Monitoring Station

Siti Aisyah Zakaria, Nor Azrita Mohd Amin, Noor Fadhilah Ahmad Radi, Nasrul Hamidin

***Abstract*—**Higher ground-level ozone (GLO) concentration adversely affects human health, vegetations as well as activities in the ecosystem. In Malaysia, most of the analysis on GLO concentration are carried out using the average value of GLO concentration, which refers to the centre of distribution to make a prediction or estimation. However, analysis which focuses on the higher value or extreme value in GLO concentration is rarely explored. Hence, the objective of this study is to classify the tail behaviour of GLO using generalized extreme value (GEV) distribution estimation the return level using the corresponding modelling (Gumbel, Weibull, and Frechet) of GEV distribution. The results show that Weibull distribution which is also known as short tail distribution and considered as having less extreme behaviour is the best-fitted distribution for four selected air monitoring stations in Peninsular Malaysia, namely Larkin, Pelabuhan Kelang, Shah Alam, and Tanjung Malim; while Gumbel distribution which is considered as a medium tail distribution is the best-fitted distribution for Nilai station. The return level of GLO concentration in Shah Alam station is comparatively higher than other stations. Overall, return levels increase with increasing return periods but the increment depends on the type of the tail of GEV distribution's tail. We conduct this study by using maximum likelihood estimation (MLE) method to estimate the parameters at four selected stations in Peninsular Malaysia. Next, the validation for the fitted block maxima series to GEV distribution is performed using probability plot, quantile plot and likelihood ratio test. Profile likelihood confidence interval is tested to verify the type of GEV distribution. These results are important as a guide for early notification on future extreme ozone events.

***Keywords*—**Extreme value theory, generalized extreme value distribution, ground-level ozone, return level.

I. INTRODUCTION

GLO is an air pollutant, and it has become an environmental issue in the last decade. These issues have happened not only in Malaysia but worldwide. High GLO concentration poses risks to human health and the environment. Some study from [1] state that a 10 ppb increase in the previous week's O_3 concentration was associated with a 0.52% increase in daily mortality and a 0.64% increase in cardiovascular and respiratory mortality while, in environmental problem, GLO is formed when two types of pollutants which are volatile organic compounds (VOCs) and nitrogen oxide (NOx) react in sunlight. The main source of ozone formation is fossil fuel combustion from industrial areas.

Continuous monitoring for the level of GLO concentration is important to ensure that the level of ozone concentration does not exceed the standard recommendation stated by the Malaysia Ambient Air Quality Standard (MAAQS) and National Ambient Air Quality Standard (NAAQS). MAAQS recommends that average ozone concentration value should not exceed 0.06 ppm in 8 hours. This continuous monitoring policy is specifically stated in the National Policy on the Environment [2] to ensure continuous improvement in the productivity and quality of the environment.

Recent studies on GLO concentration level in Peninsular Malaysia are done using time series analysis such as [3] and [4]; regression method study by [5]-[7]; as well as principal component analysis (PCA) by [7]-[9]. All of these analyses are done using classical statistics that focus on average behaviour. However, there is a lack of study using extreme value theory (EVT) which considers extreme value in the data.

EVT is the statistical modelling of extreme data found in the tails of a distribution. This method focused on the extreme value that refers to either very large (maximum) or minimal values (minimum) in a probability distribution [10]-[12]. Two primary methods in the analysis of extreme cases are by block maxima or threshold method. In this study, block maxima are used where the maximum value of observation is in length of the interval, T. This study has chosen the monthly maxima, that refer to maximum value obtained each month along the year, as the block. The limiting distribution of this approach corresponds to GEV distribution. This generalization focused on stationary case with GEV model where parameters μ, σ, ξ are constant [13]-[15].

GLO levels that are too high can be categorized as an extreme case that can be analysed using extreme value. GLO is obtained in the troposphere, the layer of atmosphere that extends from the earth's surface to about 10 miles up, which is also known as "bad" ozone. This phenomenon is also recognized as secondary air pollutant because it is produced when two primary pollutants react in sunlight and stagnant air [16]. These two primary pollutants are VOCs and NOx.

Z S Aisyah is with Institute of Engineering Mathematics, Universiti Malaysia Perlis, Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia (corresponding author, phone: +60108968643; e-mail: aisyahzakaria@unimap.edu.my).

A M N Azrita is with Institute of Engineering Mathematics, Universiti Malaysia Perlis, Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia (e-mail: norazrita@unimap.edu.my).

H Nasrul it was in Centre for Professional Development and Continuing Education (PDCE), UniMAP @Kuala Lumpur, 57000 Kuala Lumpur, Malaysia. (e-mail: nasrulhamidin@unimap.edu.my).

R A N Fadhilah it was in Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia. (noorfadhilah@tmsk.uitm.edu.my).

Excessive exposure to GLO can cause asthma, irritation of the respiratory system, permanent lung damage as well as adverse effects on vegetations and animals. People living in areas where GLO concentration levels are regularly high may find that the initial signs go away over time, but ozone continues to cause lung damage even when symptoms have disappeared [17]. The best technique to protect your health is to find out when ozone levels are raised in your area and take simple defence measures to minimize exposure to unhealthy levels of ozone, even when you do not feel obvious symptoms.

The aims of this study are to classify the tail behaviour of GLO using GEV distribution, which is either Gumbel, Weibull or Fréchet distribution; secondly, to estimate the return level using the corresponding distribution (Gumbel, Weibull or Fréchet) type of GEV distribution.

II. METHODOLOGY

The strategy employed to classify the tail distribution in this study does not interest the researcher's attention. There is a greater emphasis on the distribution in the middle. As a result, this research proposes a GEV distribution to categorize data that focuses on the extreme component of the distribution, also known as tailed distribution.

A. GEV Distribution

GEV is a family of asymptotic distributions that describe the behaviour of extreme conditions. GEV distribution consists of Gumbel, Fréchet, and Weibull families which are known as type I, II and III extreme value distributions [18]. The cumulative distribution function (cdf) for GEV is given in (1):

$$G(\mu,\sigma,\xi) = \begin{cases} exp\left\{-\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\}, & \xi \neq 0 \\ exp\left\{-exp\left[-\left(\dfrac{x-\mu}{\sigma}\right)\right]\right\}, & \xi = 0. \end{cases} \tag{1}$$

Each family of GEV distribution has a location (μ), scale (σ) and shape (ξ) parameter with $\mu = \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$, respectively. The type of distribution is inferred by the shape parameter (ξ); which describes the tail behaviour of the data distribution. Specifically, when $\xi = 0$, the GEV distribution is Gumbel distribution which supports $\mathbb{R}$ with a light tail decaying exponentially; when $\xi > 0$, it refers to Fréchet distribution which supports $x \geq \mu + \frac{\sigma}{\xi}$ with a heavy tail decaying polynomially, and when $\xi < 0$ it corresponds to bounded tail (Weibull distribution) which supports $x \leq \mu - \frac{\sigma}{\xi}$. These three-distributions indicate whether there is more or less extreme behaviour in the data analysed. This method has been done in [19]-[26].

B. Selected Period

The selection period for this study depends on blocks with the same length *n* and the maximum of each block form a series of block maxima which are fitted to GEV distribution. The selection of block size is important as if the block size is too small, the result can lead to bias and if they are too large, the number of blocks maxima generated are too small, leading to large estimation variance and length of period [18].

In this study, GLO data are blocked into monthly maxima for 17 years (2000 until 2016) which are complete data extract that available to us for analysis part. The maximum value for each month is chosen as extreme monthly maxima for analysis. The modelling of GEV distribution needs the assumption of X to have an independent random variable with common distribution, F.

Data used are from continuous ambient air quality monitoring stations (CAQMS) in Peninsular Malaysia which are owned by the Department of Environment, Malaysia (DOE) but are managed and operated by a private company, Alam Sekitar Malaysia Sdn. Bhd. (ASMA). The ozone concentration in all of the stations is measured using Teledyne Ozone Analyzer Model 400E UV Absorption. The analyser uses a system based on the Beer-Lambert law for measuring low ranges of ozone in ambient air.

C. Classification Approach

In this work, the classification consists of several procedures. It involves the parameter estimates by using the MLE which became a standard benchmark in this field due to estimation of the model parameter. The straight MLE approach works reasonably well for this GLO data. Based on [18], under the assumption that $Z_1,.....,Z_m$ is the independent variable with GEV distribution, the log-likelihood function for the GEV parameters are in (2) and (3):

$$\begin{aligned} l\left(\mu,\sigma,\xi\right) = & -mlog\sigma \\ & -\left(1+\frac{1}{\xi}\right)\sum_{i=1}^{m} log\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] \\ & -\sum_{i=1}^{m}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right], \quad \xi \neq 0 \end{aligned} \tag{2}$$

$$\begin{aligned} l\left(\mu,\sigma,\xi\right) = & -mlog\sigma - \sum_{i=1}^{m}\left(\frac{x_i-\mu}{\sigma}\right) \\ & -\sum_{i=1}^{m} exp\left[-\left(\frac{x_i-\mu}{\sigma}\right)\right], \quad \xi = 0 \end{aligned} \tag{3}$$

Although other parameter estimations might be better than MLE, MLE has the benefit of adaptability to changes in the model structure. In addition, MLE can also employ censored information without difficult [27].

In order to evaluate the correct model, tests are carried out using suitable tests such as likelihood ratio (LR) test and profile likelihood confidence interval. LR test is used to compare the fit of the models where the hypothesis is $H_0 : \xi = 0$ (Gumbel distribution) and alternative hypothesis is $H_1 : \xi < 0$ (Weibull distribution). The best model is determined by deriving the probability or *p*-value of the difference in LR test. The null hypothesis is rejected in favour of the alternative of significance if *p*-value $< \alpha$, where α is the level of significance (0.05). The idea of profile likelihood confidence interval is to invert a likelihood-ratio test to obtain a confidence interval for the parameter in question. Meanwhile, the diagnostics plots such as quantile plot, density plot and return level plot are used. A quantile plot that greatly deviates from a straight line suggests that model assumptions may be invalid for the data plotted. For density plot, a model line is compared against the empirical data line. And lastly, return level plot shows the return period against the return level, with an estimated 95% confidence interval. The best-fitted model has points on the probability plot that lie on the unit diagonal, but these graphical tests are only used as complements to a statistical test. According to [27], statistical analysis is more accurate than the graphical test.

In extreme investigations, this method involves return levels, which means the value of extreme cases that occur, on average, once in a given period whether in a year, month, or others. The information from return levels could be of benefit to DOE, to monitor ozone concentration level in all stations and can be used to take action before something happens [10]. Z_p in (4) defines the return level for GEV distribution with return period $\frac{1}{p}$.

$$Z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left\{1 - \left[-\log\left(1-p\right)\right]^{-\xi}\right\}, & when\ \ \xi \neq 0 \\ \mu - \sigma\log\left(1-p\right), & when\ \ \xi = 0. \end{cases} \tag{4}$$

Based on the analysis of the goal, this study used extreme GLO in Peninsular Malaysia where 26 stations are located all over the country. From these 26 stations, only five stations are chosen based on the top five highest percentage of ozone concentration exceeding the Malaysian Ambient Air Quality Standard MAAQS of 0.06 ppm which used as a reference. These five selected stations are categorized as industrial, urban, and suburban as in Table I and Fig. 1.

The first station, Larkin, is in the south of Johor with a total population of more than 1.3 million (2010) [1] and is close to the largest industrial area in Johor. Meanwhile Nilai station is located close to the residential and industrial city with a total population of more than 1.3 million [1]. The stations in Shah Alam are also located at high-density residential areas, which are frequently affected by traffic-related pollution and are closed to major industrial areas. The next station is Pelabuhan Kelang which is close to the main road at Persiaran Raja Muda Musa with residential areas surrounding it. The population density estimates here is around 200 thousand people [1] with a variety of activities that contribute to the sources of ozone formation which affect ozone concentration. The last station, Tanjung Malim, is located at Universiti Pendidikan Sultan Idris (UPSI), close to residential areas and some small industrial activities.

TABLE I
LOCATIONS AND DESCRIPTIONS OF SELECTED AIR MONITORING STATIONS IN PENINSULAR MALAYSIA

Areas	Air monitoring stations	Coordinates	Strata
Larkin, Johor	Institut Perguruan Malaysia, Temenggong Ibrahim, Larkin, Johor Bharu	N 1.491781 E 103.738173	Industrial
Nilai, Negeri Sembilan	Tmn. Semarak (Phase II), Nilai	N 2.811261 E 101.799275	Industrial
Shah Alam, Selangor	Sek. Keb. TTDI Jaya, Shah Alam	N 3.106797 E 101.557303	Urban
Pelabuhan Kelang, Selangor	Sek. Men. (P) Raja Zarina, Kelang	N 3.011732 E 101.409796	Urban
Tanjung Malim, Perak	Universiti Pendidikan Sultan Idris, Tanjung Malim	N 3.685105 E 101.524086	Suburban

Fig. 1 Map of five selected air monitoring stations

III. CLASSIFICATION ASSESSMENT

A. Descriptive Statistics

Table II shows the descriptive statistics for monthly block maxima in five selected stations. The unit of measurement is particle per million (ppm). It shows that the maximum concentrations for the five stations have exceeded the Malaysian Ambient Air Quality Guidelines (MAAQG) level for the hourly average of 0.06 ppm with the highest maximum concentration in Larkin, Johor, 0.20 ppm. All of the data in the five stations are moderately skewed with the range of skew value around -1 and 1 [28]. Three stations, namely Nilai, Larkin and Tanjung Malim are positive skewed, that is the right tail of distribution is relatively longer than the left side, denoting that most of the data are concentrated on the left of the mean. The other two stations, Shah Alam and Pelabuhan Kelang are negatively skewed that is more data are concentrated to the right of the mean. Even though Larkin has the highest maximum reading of ozone concentration, Shah Alam area has the highest percentage of the number of block maximum with 71.57% 0.01 ppm MAAQG value compared to other stations. This is because Shah Alam is located in an industrial and urban area with a larger population and more industrial activities as well.

TABLE II
DESCRIPTIVE STATISTICS OF GLO CONCENTRATION IN FIVE SELECTED STATIONS

Station	Strata	Mean	Max	SD	Skew	Exc (%)
Nilai	Industrial	0.09	0.15	0.02	0.79	24.51
Larkin	Industrial	0.08	0.20	0.02	0.69	17.65
Shah Alam	Urban	0.11	0.17	0.03	-0.50	71.57
Pelabuhan Kelang,	Urban	0.08	0.13	0.02	-0.07	17.16
Tanjung Malim	Suburban	0.09	0.16	0.02	0.12	37.75

[a]N= 204 Sample Size, SD = Standard Deviation, Exc = number of block maximum, which exceeds the 0.06 ppm MAAQG value.

B. Parameter Estimates

The finding for the first objective in Table III is obtained by assuming that the parameter estimates are stationary models and it is shown that there is no trend in the data. In a stationary model case, all parameters are considered as a constant which is time independent. MLE is used to estimate all parameters.

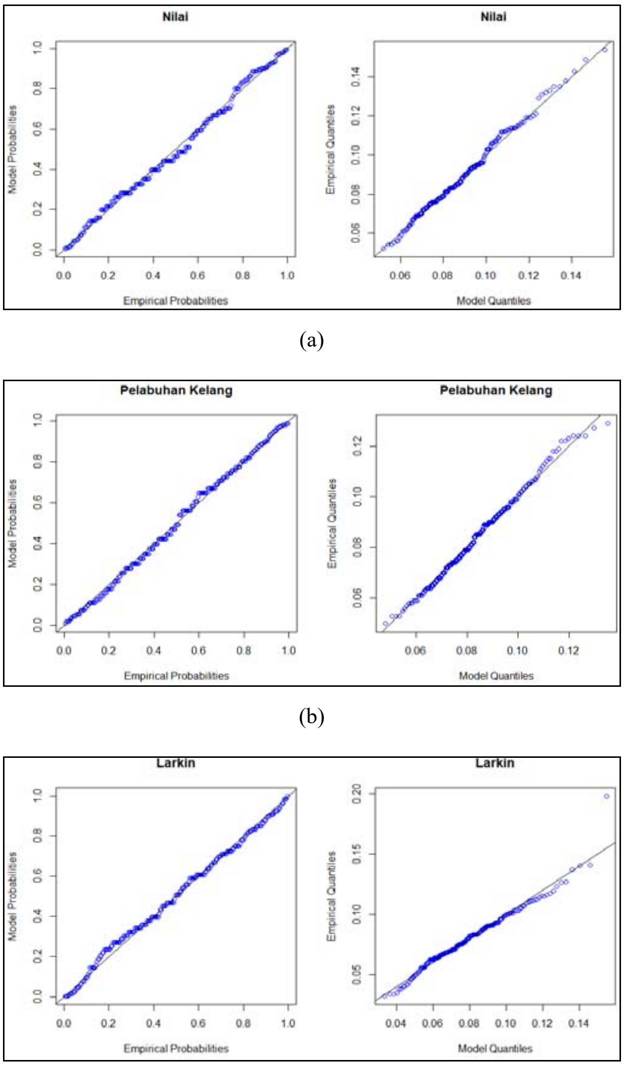
The negative values of shape parameter (ξ) in Table III for all stations indicate that the distribution is bounded on the upper end (or short-tailed). This distribution is categorized as Type III or Weibull distribution. But, for Nilai station, the shape value is so small that it can be categorized as Type I or Gumbel distribution, and this will be verified approve by testing using LR test and profile likelihood confidence interval.

TABLE III
PARAMETER ESTIMATES FOR GEV DISTRIBUTION

Station	Strata	μ (SE)	σ (SE)	ξ (SE)
Nilai	Industrial	0.0798(0.0013)	0.0160(0.0009)	-0.0457(0.0508)
Larkin	Industrial	0.0693(0.0017)	0.0222(0.0011)	-0.1291(0.0213)
Shah Alam	Urban	0.1047(0.0020)	0.0271(0.0013)	-0.3509(0.0314)
Pelabuhan Kelang	Urban	0.0773(0.0014)	0.0188(0.0009)	-0.2974(0.0348)
Tanjung Malim	Suburban	0.0784(0.0018)	0.0237(0.0013)	-0.2202(0.0414)

C. Model Diagnostics

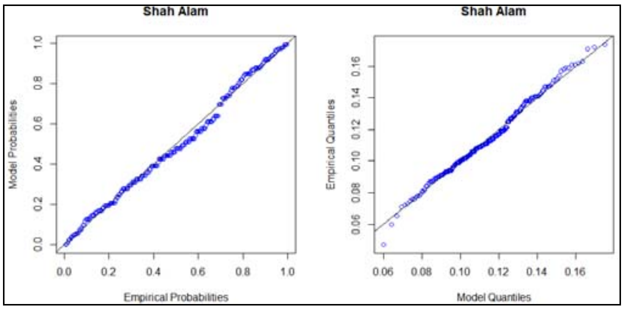
Results from parameter estimation have shown that all of the distributions fit well with GEV distribution and the shape parameter for all five stations having negative value. Based on these results, Weibull distribution is the best distribution that fits all the selected stations. LR test is used to confirm the shape parameter by performing the null hypothesis $H_0 : \xi = 0$ (Gumbel distribution), and it is rejected. Furthermore, Profile Likelihood Confidence Interval test is also carried out to show that all shape parameters are positioned within the negative interval. By referring to Table IV, the model diagnostics or model checking for the five selected stations has shown that only one station, Nilai, has a p-value of more than 0.05 which is not significant. Hence, the Gumbel distribution is the best distribution to describe the extreme GLO for Nilai station. This decision is proven by conducting a second test with the result (-0.1385, 0.01) indicating that the value of GLO concentration in Nilai can be within -0.1385 and 0.01 that can contribute to Gumbel distribution with $\xi = 0$. The remaining four stations have a p-value of less than 0.05 that fits Weibull distribution, and this is proven by the second test. Besides, graphical inspection using probability and quantile plots has shown that all points are scattered near to the linear line. Thus, there is no doubt on the validity of the fitted model (Fig. 2). The return level curve asymptotes to a finite level because of the negative estimate of ξ as in Fig. 3, however, since the estimate is close to zero, the estimated curve is close to linear. Lastly, all three diagnostic plots lend support to the fitted GEV model.



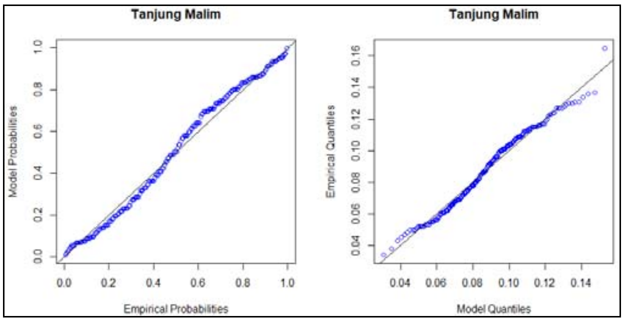
(a)

(b)

(c)

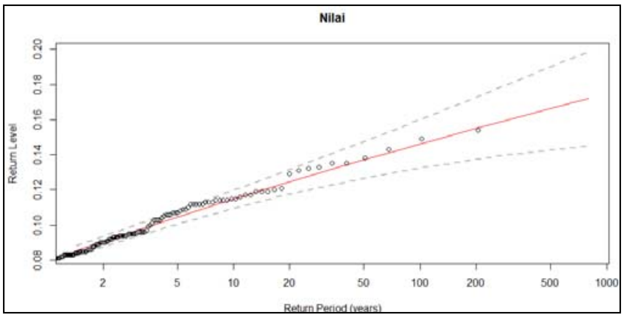


(d)

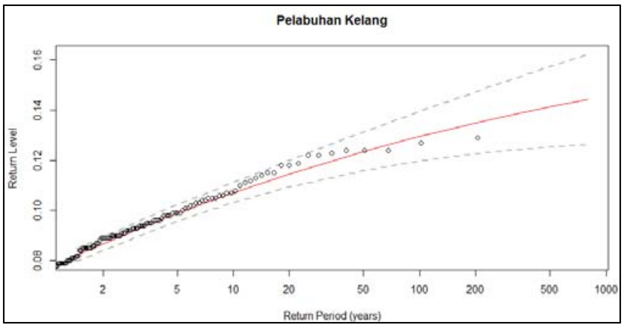


(e)

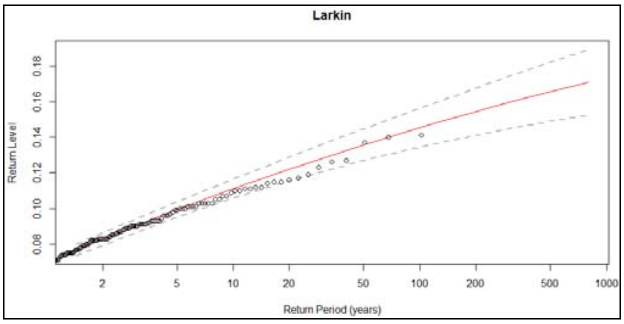
Fig. 2 Probability plot and quantile plot for (a) Nilai, (b) Pelabuhan Kelang, (c) Larkin, (d) Shah Alam, and (e) Tanjung Malim Stations



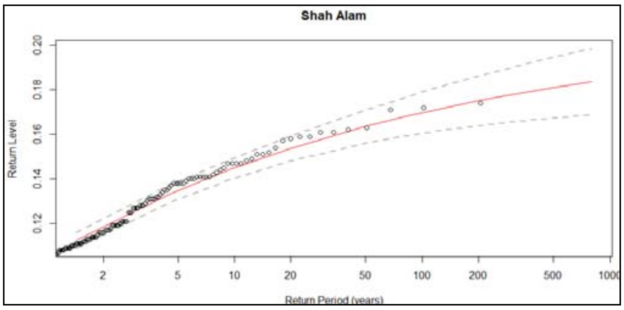
(a)



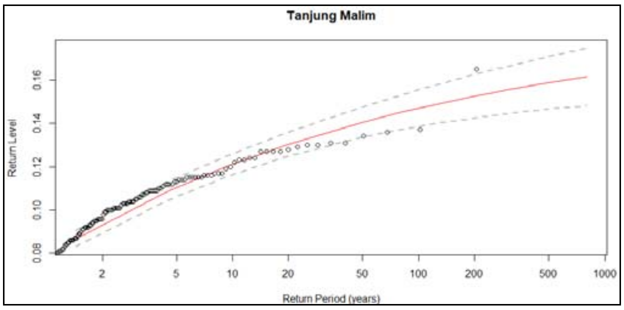
(b)



(c)



(d)



(e)

Fig. 3 Return level plot for (a) Nilai, (b) Pelabuhan Kelang, (c) Larkin, (d) Shah Alam, and (e) Tanjung Malim Stations

TABLE IV
LR TEST AND PROFILE LIKELIHOOD CONFIDENCE INTERVAL (CI)

Station	Strata	Statistics Test	
		LR (p-value < 0.05)	Profile likelihood CI
Nilai	Industrial	0.3842	(-0.1385, 0.0100)
Larkin	Industrial	0.0000	(-0.1646, -0.0802)
Shah Alam	Urban	0.0000	(-0.4000, -0.2895)
Pelabuhan Kelang	Urban	0.0000	(-0.3643, -0.2303)
Tanjung Malim	Suburban	0.0002	(-0.2739, -0.1212)

D. Return Level Estimates

Table V shows the results of return level for three return periods (2-month, 10-month, and 20-month). Based on Fig. 3, the return level plot indicates that the dotted line is linear up to 20-month of return period which means that to get accurate estimation researchers can make estimation around this return value, however, after that the estimation is not valuable. So, this study has estimated the value according to the three return periods, as shown in Table V. The results show that GLO concentration increases with the increase in return period for all stations. And all readings exceed the MAAQG guideline (0.06 ppm). The return levels of GLO concentration in Shah Alam station are comparatively higher than other stations. This estimation can be used as a guide to formulate solutions to handle extreme ozone events which in turn would prevent the problem of air pollution from worsening in the future. Amin and Zakaria [13] state that the return level explains the maximum level that can be estimated in average for the particular years whereas by using the forecasting value for average data, it reflects the predicted values for the next particular years period.

TABLE V
RETURN LEVEL OF MONTHLY MAXIMA GLO CONCENTRATION DATA

Stations	Strata	Return Period		
		2-month	10-month	20-month
Nilai	Industrial	0.0856	0.1140	0.1242
		(0.0829, 0.0883)	(0.1090, 0.1189)	(0.1174, 0.1310)
Larkin	Industrial	0.0772	0.1126	0.1240
		(0.0738, 0.0806)	(0.1074, 0.1177)	(0.1180, 0.1300)
Shah Alam	Urban	0.1141	0.1469	0.1548
		(0.1102, 0.1179)	(0.1431, 0.1507)	(0.1508, 0.1587)
Pelabuhan Kelang	Urban	0.0838	0.1081	0.1144
		(0.0810, 0.0866)	(0.1051, 0.1112)	(0.1110, 0.1177)
Tanjung Malim	Suburban	0.0868	0.1205	0.1301
		(0.0830, 0.0905)	(0.1157, 0.1252)	(0.1246, 0.1355)

IV. CONCLUSION

In conclusion, this study has shown that Weibull distribution is the best-fitted distribution for the four selected stations which are Pelabuhan Kelang, Larkin, Shah Alam and Tanjung Malim stations. This distribution, also known as bounded tail, indicates less extreme behaviour. Otherwise, Gumbel distribution is the best-fitted distribution for Nilai station. This distribution is also recognized as having less extreme behaviour. Less extreme here means that the level of GLO concentration in these five stations are considered as moderate and still in healthy conditions. Inspection of the probability and quantile plots has shown points that are scattered near to the linear line. Thus, there is no doubt on the validity of the fitted model. All diagnostic plots lend support to the fitted GEV model. Lastly, it is found that the return levels of GLO concentration in Shah Alam station are comparatively higher than other stations. Overall, return levels increase with increasing return period but the increasing pattern is dependent on the type of the tail of GEV distribution. These results can be used as a guide for early detection in the effort to overcome future air pollution due to ozone pollutants.

ACKNOWLEDGMENT

We would like to thank the Department of Environment (DoE), Malaysia, for supplying the ground-level ozone data specifically in Peninsular Malaysia, which are used to carry out this study. This study is also supported by FRGS grant with the code FRGS/1/2019/STG06/UNIMAP/02/4.

REFERENCES

[1] M. A. B. A. Tajudin *et al.*, "Risk of concentrations of major air pollutants on the prevalence of cardiovascular and respiratory diseases in urbanized area of Kuala Lumpur, Malaysia," *Ecotoxicol. Environ. Saf.*, vol. 171, no. December 2018, pp. 290–300, 2019, doi: 10.1016/j.ecoenv.2018.12.057.

[2] M. Department of Environment (DoE), "National Policy on the Environment." Ministry of Science, Technology and the Environment, Malaysia, pp. 1–24, 2002.

[3] M. Ismail, "Time-series analysis of ground-level ozone in Muda Irrigation Scheme Area (MADA), Kedah," *J. Sustain. Sci. Manag.*, vol. 6, no. 1, pp. 79–88, 2011.

[4] N. R. Awang, N. A. Ramli, N. I. Mohammed, and A. S. Yahaya, "Time Series Evaluation of Ozone Concentrations in Malaysia Based on Location of Monitoring Stations Time Series Evaluation of Ozone Concentrations in Malaysia Based on Location of Monitoring Stations," no. June, 2016.

[5] Mohamed Elnour Yassen, Jamaluddin Md. Jahi, and Shaharuddin Ahmad, "Evaluation of long term trends in oxide of nitrogen concentrations in the Klang Valley region, Malaysia," *Malaysian J. Environ. Manag.*, vol. 6, pp. 59–72, 2005, (Online). Available: http://www.ems-malaysia.org/mjem/index.html%5Cnhttp://journalarticle.ukm.my/2229/%5Cnhttp://journalarticle.ukm.my/2229/1/2005_4_ElNour.pdf.

[6] N. I. Mohammed, N. Azam Ramli, A. Shukri Yahya, and N. Adyani Ghazali, "Aotx Indices in Response to the Crops Yield Under the Malaysia Climate," *Ijrras*, vol. 7, no. 3, pp. 304–309, 2011.

[7] J. M. Rajab, M. Z. MatJafri, and H. S. Lim, "Combining multiple regression and principal component analysis for accurate predictions for column ozone in Peninsular Malaysia," *Atmos. Environ.*, vol. 71, pp. 36–43, 2013, doi: 10.1016/j.atmosenv.2013.01.019.

[8] K. C. Tan, H. S. Lim, and M. Z. M. Jafri, "Analysis of total column ozone in Peninsular Malaysia retrieved from SCIAMACHY," *Atmos. Pollut. Res.*, vol. 5, no. 1, pp. 42–51, 2014, doi: 10.5094/APR.2014.006.

[9] M. T. Latif, E. Z. Abidin, and S. M. Praveena, "The Assessment of Ambient Air Pollution Trend in Klang Valley," *World Environ.*, vol. 5, no. 1, pp. 1–11, 2015, doi: 10.5923/j.env.20150501.01.

[10] A. Z. Mohd Amin, N. A., Adam, M. B., Aris, "Extreme Value Analysis for Modeling High PM10 Level in Johor Bahru," *J. Teknol.*, vol. 1, no. 76, pp. 171–179, 2015.

[11] H. Hasan, N. Fadhilah, A. Radi, and S. Kassim, "Modeling the Distribution of Extreme Share Return in Malaysia Using Generalized Extreme Value (GEV) Distribution," vol. 89, pp. 82–89, 2012, doi: 10.1063/1.4724121.

[12] H. Hassan, N. Salam, and M. D. Adam, "Modelling Extreme Temperature in Malaysia Using Generalized Extreme Value Distribution," *Int. J. Math. Comput. Phys. Electr. Comput. Eng.*, vol. 7, no. 6, pp. 983–989, 2013.

[13] M. Amin, N. Azrita, and Z. S. Aisyah, "Comparison on the Analysis on PM10 Data based on Average and Extreme Series," vol. 05025, pp. 1–5, 2018.

[14] N. S. Muhammad, A. I. Akashah, and J. Abdullah, "Analysis of extreme rainfall indices in peninsular Malaysia," *J. Teknol.*, vol. 78, no. 9–4, pp. 15–20, 2016, doi: 10.11113/jt.v78.9677.

[15] S. El Adlouni, T. B. M. J. Ouarda, X. Zhang, R. Roy, and B. Bobe, "Generalized maximum likelihood estimators for the nonstationary generalized extreme value model," vol. 43, pp. 1–13, 2007, doi: 10.1029/2005WR004545.

[16] Tor H. Oiamo, "Ground-Level Ozone," in *Encyclopedia of Quality of Life and Well-Being Research*, Springer Netherlands, 2014, pp. 2619–2619.

[17] J. M. Rajab, H. S. Lim, and M. Z. Matjafri, "Monthly distribution of diurnal total column ozone based on the 2011 satellite data in Peninsular Malaysia," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 16, no. 1, pp. 103–109, 2013, doi: 10.1016/j.ejrs.2013.04.003.

[18] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London, 2001.

[19] S. Hazarika, P. Borah, and A. Prakash, "The assessment of return probability of maximum ozone concentrations in an urban environment of Delhi: A Generalized Extreme Value analysis approach," *Atmos. Environ.*, vol. 202, no. August 2018, pp. 53–63, 2019, doi:

10.1016/j.atmosenv.2019.01.021.

[20] F. Calderón-Vega, C. Mösso, A. D. García-Soto, and E. Delgadillo-Ruiz, “Single Site Extreme Wave Analysis in the Pacific Ocean Comparing Stationary and Non-stationary GEV Models,” *Curr. J. Appl. Sci. Technol.*, vol. 32, no. 6, pp. 1–12, 2019, doi: 10.9734/cjast/2019/v32i630038.

[21] H. Hasan, N. F. A. Radi, and S. Kassim, “Modeling the distribution of extreme share return in Malaysia using Generalized Extreme Value (GEV) distribution,” in *The 5th International Conference on Research and Education in Mathematics*, 2012, vol. 1450, no. June, pp. 82–89, doi: 10.1063/1.4724121.

[22] A. J. Cannon, “Computers & Geosciences GEVcdn: An R package for nonstationary extreme value analysis by generalized extreme value conditional density estimation network $,” *Comput. Geosci.*, vol. 37, no. 9, pp. 1532–1533, 2011, doi: 10.1016/j.cageo.2011.03.005.

[23] F. De Paola, M. Giugni, F. Pugliese, A. Annis, and F. Nardi, “GEV parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in African test cities,” *Hydrology*, vol. 5, no. 2, pp. 1–22, 2018, doi: 10.3390/hydrology5020028.

[24] S. Jiang and L. Kang, “Flood frequency analysis for annual maximum streamflow using a non-stationary GEV model,” *E3S Web Conf.*, vol. 79, 2019, doi: 10.1051/e3sconf/20197903022.

[25] Anita R., “Identification of Climate Change with Generalized Extreme Value (GEV) Distribution Approach,” *J. Phys. Conf. Ser.*, vol. 423, 2013, doi: 10.1088/1742-6596/423/1/012026.

[26] H. Hasan, N. Fadhilah, B. A. Radi, and S. B. Kassim, “Modeling of Extreme Temperature Using Generalized Extreme Value (GEV) Distribution: A Case Study of Penang,” in *Proceedings of the world Congress on Engineering*, 2012, vol. I, pp. 4–9.

[27] H. Hasan, N. Salam, and S. Kassim, “Modeling annual extreme temperature using generalized extreme value distribution: A case study in Malaysia,” *AIP Conf. Proc.*, vol. 1522, pp. 1195–1203, 2013, doi: 10.1063/1.4801267.

[28] K. Pearson, “Skew Variation in Homogeneous Material,” *Philos. Trans. R. Soc. London. A*, vol. 186, no. 1895, pp. 343–414, 1895, (Online). Available: http://www.jstor.org/stable/90649%0Ahttp://www.jstor.org/page/info/about/policies/terms.jsp%0Ahttp://www.jstor.org.